

COMPUTER AIDED METHODS FOR LOWER BOUNDS ON THE BORDER RANK

JONATHAN D. HAUENSTEIN, CHRISTIAN IKENMEYER, AND J.M. LANDSBERG

ABSTRACT. We present new methods for determining polynomials in the ideal of the variety of bilinear maps of border rank at most r . We apply these methods to several cases including the case $r = 6$ in the space of bilinear maps $\mathbb{C}^4 \times \mathbb{C}^4 \rightarrow \mathbb{C}^4$. This space of bilinear maps includes the matrix multiplication operator M_2 for two by two matrices. We show these newly obtained polynomials do not vanish on the matrix multiplication operator M_2 , which gives a new proof that the border rank of the multiplication of 2×2 matrices is seven. Other examples are considered along with an explanation of how to implement the methods.

Acknowledgements. We thank Peter Bürgisser for important discussions and suggestions.

1. INTRODUCTION

Lower bounds in complexity theory are considered difficult to obtain. We describe a new method for obtaining lower bounds on the *border rank* which is based on a new way to find polynomials that vanish on bilinear maps $T : \mathbb{C}^{\mathbf{a}} \times \mathbb{C}^{\mathbf{b}} \rightarrow \mathbb{C}^{\mathbf{c}}$ of low border rank.

1.1. Rank and border rank. Let $\mathbb{C}^{\mathbf{a}*} := \{f : \mathbb{C}^{\mathbf{a}} \rightarrow \mathbb{C} \mid f \text{ is linear}\}$ denote the dual vector space to $\mathbb{C}^{\mathbf{a}}$. That is, if an element of $\mathbb{C}^{\mathbf{a}}$ is represented by a column vector of height \mathbf{a} , then $\mathbb{C}^{\mathbf{a}*}$ corresponds to row vectors, and the evaluation is just row-column matrix multiplication. A bilinear map $T : \mathbb{C}^{\mathbf{a}} \times \mathbb{C}^{\mathbf{b}} \rightarrow \mathbb{C}^{\mathbf{c}}$ has *rank one* if there exist $\alpha \in \mathbb{C}^{\mathbf{a}*}$, $\beta \in \mathbb{C}^{\mathbf{b}*}$, and $c \in \mathbb{C}^{\mathbf{c}}$ such that $T(a, b) = \alpha(a)\beta(b)c$. The rank one bilinear maps are in some sense the simplest bilinear maps, and T is said to have *rank r* if r is the minimum number of rank one bilinear maps which sum to T . This r is sometimes called the *tensor rank* of T . If one views multiplication by constants as a “free” operation, then the rank differs at most by a factor of two from the minimal number of multiplications of variables that is needed to compute T , see [3, Ch. 14] for more information.

Since the set of all bilinear maps $\mathbb{C}^{\mathbf{a}} \times \mathbb{C}^{\mathbf{b}} \rightarrow \mathbb{C}^{\mathbf{c}}$ is a vector space of dimension \mathbf{abc} , it is natural to talk about polynomials on the space of bilinear maps $\mathbb{C}^{\mathbf{a}} \times \mathbb{C}^{\mathbf{b}} \rightarrow \mathbb{C}^{\mathbf{c}}$. Unfortunately, one cannot test directly for the tensor rank by the vanishing of polynomials, since the common zero locus of the set of all polynomials vanishing on the set of bilinear maps of rank at most r is, typically, larger than the set of bilinear maps of rank at most r . This may be described precisely using the language of algebraic geometry: for the purposes of this article, we define an *algebraic variety* (or simply a *variety*) to be the common zero locus of a collection of polynomials that is *irreducible*, in the sense that it cannot be written as a union of two zero loci. A (proper) *Zariski closed subset* of a variety X is the common zero locus of a collection of polynomials restricted to X , and a *Zariski open subset* is the complement of a Zariski closed set. The set of bilinear maps of rank r is a Zariski open subset of the algebraic variety formed from the zero set of all

Key words and phrases. border rank, matrix multiplication, MSC 68Q17.

Research of Hauenstein supported in part by AFOSR grant FA8650-13-1-7317 and NSF grant DMS-1262428. Research of Ikenmeyer supported in part by DFG grant BU 1371/3-2. Research of Landsberg supported in part by NSF grant DMS-1006353.

the polynomials having the set of bilinear maps of rank at most r in their zero set. The *border rank* of a tensor T is defined to be the smallest r such that all polynomials vanishing on the set of bilinear maps of rank at most r also vanish at T , and one writes $\mathbf{R}(T) = r$. In this case, T is arbitrarily close, in any measure, to a bilinear map of rank r (including the possibility that the rank of T is r). We let $\sigma_{r;\mathbf{a},\mathbf{b},\mathbf{c}}$ denote the set of bilinear maps of border rank at most r . It is an algebraic variety. When $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are understood from the context, we simply write σ_r .

1.2. Results. We introduce a new technique based on machine learning that finds, with high probability, where equations which vanish on the variety of bilinear maps of border rank at most r can be found. Once one knows where to look, we can use methods which began in [16] and were refined in [1, 5] to find the actual equations. The vanishing of these equations can be sometimes proved rigorously and other times proven with extremely high probability depending on the application. With these equations and given a tensor T , one can then show its border rank is greater than r if such a polynomial P satisfies $P(T) \neq 0$. Of special interest in this paper will be border rank of the matrix multiplication tensor

$$M_2 := \sum_{i,j,k=1}^2 e_{i,j} \otimes e_{j,k} \otimes e_{k,i} \in \mathbb{C}^4 \otimes \mathbb{C}^4 \otimes \mathbb{C}^4,$$

where $(e_{i,j})$ is the standard basis of $\mathbb{C}^{2 \times 2} = \mathbb{C}^4$.

The technique we present can be applied to other *implicitization problems*. That is, we want to consider a variety

$$(1) \quad X := \overline{g(Y)}$$

where Y is (possibly a Zariski open set of) a variety and g is a system of rational functions defined on Y . In the bilinear case, the tensors of rank at most r is a dense subset of the algebraic set of tensors of border rank at most r where each tensor of rank at most r can be written as a sum of r tensors of rank at most one. In this case, Y is simply the product of r copies of the variety of tensors of rank at most one and g is the linear map corresponding to taking the sum. Another specific application arising in physics is the analysis of vacuum moduli space in (supersymmetric) field theories [8] which arise as the closure of the image under a polynomial map of an algebraic set.

Before summarizing the main results, we note that a result “with extremely high probability” means the result was proven using numerical techniques. When we write “equations,” we mean that we have written down the explicit equation. The main results presented here using our new technique are as follows.

- No polynomial of degree less than 19 vanishes on $\sigma_{6;4,4,4}$.
- The degree of $\sigma_{6;4,4,4}$, with extremely high probability, is 15,456.
- There is a 64 dimensional space of degree 19 equations that vanish on $\sigma_{6;4,4,4}$ with extremely high probability. One such equation does not vanish on the matrix multiplication tensor M_2 , giving a new, short proof with extremely high probability that the border rank of M_2 is seven.
- We construct an equation of degree 20 that vanishes on $\sigma_{6;4,4,4}$. This equation does not vanish on M_2 , giving a new, short proof that the border rank of M_2 is seven.
- The degree of the codimension six variety $\sigma_{6;3,4,6}$, with extremely high probability, is 206,472 and its ideal is empty in degree 14.
- The degree of the codimension two variety $\sigma_{7;4,4,5}$, with extremely high probability, is 44,000 and its ideal is empty in degree 56.
- The degree of the hypersurface $\sigma_{8;3,5,7}$ is 105, with extremely high probability.

- The degree of the hypersurface $\sigma_{18;7,7,7}$, with extremely high probability, is at least 187,000.

1.3. Other methods for finding equations. Very little is known about the equations of σ_r in general. One can reduce to the case of $\mathbf{a} = \mathbf{b} = \mathbf{c} = r$ via a process called *inheritance*. Additionally, there is a systematic way to determine the equations in any given degree using *multi-prolongation*. For a discussion on inheritance and multi-prolongation, see [15, §3.7]. Even though multi-prolongation is systematic, it is very difficult to utilize except in very small cases. Most known equations have been found by reducing multi-linear algebra to linear algebra. See [13, 17] for the most recent equations that go up to $\sigma_{2m-2;m,m,m}$. Some information about the ideal of $\sigma_{r;r,r,r}$ can be found using representation theory (via the algebraic Peter-Weyl Theorem) as this case is an orbit closure, see [4] for an exposition. By inheritance one could deduce the $\sigma_{r;m,n,p}$ case for any m, n, p from the $\sigma_{r;r,r,r}$ case.

1.4. Polynomials on vector spaces. We write $I(\sigma_r)$ for the set of all polynomials vanishing on σ_r , which forms an *ideal*. Since σ_r is invariant under re-scaling, we may restrict our attention to homogeneous polynomials since, in this case, a polynomial will be in the ideal if and only if all of its homogeneous components are in the ideal.

Let V be a vector space. A subset $X \subset V$ is called an *algebraic set* if it is the common zero locus of a collection of polynomials on V . Recall that we say that an irreducible algebraic set is a *variety*. If $X \subset V$ is a variety that is invariant under re-scaling, let $S^d V^*$ be the space of homogeneous polynomials of degree d on V and $I_d(X) \subset S^d V^*$ be the component of the ideal of X in degree d .

Roughly speaking (see §2 for more details), our technique for studying the equations that vanish on a variety X of positive dimension is by applying numerical algebraic geometry and machine learning techniques to finite subsets of X which lie in a common linear space. That is, we aim to study finite subsets of algebraic sets of the form $Y = X \cap \mathcal{L} \subset \mathcal{L}$ where \mathcal{L} is a general linear space of codimension at most $\dim X$. If the codimension of \mathcal{L} is $\dim X$, then Y consists of $\deg X$ points. If the codimension of \mathcal{L} is strictly less than $\dim X$, then Y is also a variety with the same degree as X . Moreover, if one considers $X \subset V$ and $Y \subset \mathcal{L}$, and defines d_X and d_Y to be the minimal degree of the nonzero polynomials in $I(X)$ and $I(Y)$, respectively, then $d_X \geq d_Y$. In particular, $\dim I_d(Y) \geq \dim I_d(X)$ for any $d \leq d_X$ with similar bounds for all $d \geq 0$ that can be developed from the corresponding Hilbert functions. Once we have inferred information about polynomials in $I(Y)$, we use representation theory to identify which modules could appear. Finally, sample vectors from these modules are used to test if the entire module is in the ideal $I(X)$ or not.

2. DECIDING WHERE TO GO HUNTING

The basic idea of our algorithm is to combine the ability of numerical algebraic geometry to compute points on certain subsets of a variety with methods in machine learning to obtain information about this subset from the computed points. We first describe needed concepts from numerical algebraic geometry and then from machine learning.

At a basic level, the algorithms of numerical algebraic geometry (see [21] for general background information) perform numerical computations on varieties where each variety is represented by a data structure called a *witness set*. Let f be a polynomial system. The common zero locus of f is an algebraic set that can be decomposed uniquely into finitely many varieties, none of which is contained in the union of the others. If X is one of these varieties, called an *irreducible component* of the zero locus of f , then a witness set for X is the triple $\{f, L, W\}$ where the zero set of L defines a general linear subspace of codimension equal to the dimension of X and W is the intersection of X with this linear subspace defined by L . Given one point in W ,

arbitrarily many points on X can be computed in a process called *sampling*. In numerical terms, computing a point “on” a variety means that we have both a numerical approximation of the point along with an algorithm that can be used to approximate the point to arbitrary accuracy.

This witness set description is not useful for the problems at hand since, for each of the varieties X under consideration, we do not assume that we have access to a polynomial system f let alone *any* nonzero polynomials which vanish on X . However, we do assume that we have a description of X in the form (1). In fact, by adding variables and clearing denominators, we can assume that $X := \pi(\overline{Z})$ where π is a projection map and Z is an irreducible component of the zero locus for some polynomial system F . This is demonstrated in the following simple example.

Example 2.0.1. The set $X := \{(x, y) \in \mathbb{C}^2 \mid x^2 + y^2 = 1\}$ is equal to $\overline{g(Y)}$ where

$$g(t) = \left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2} \right) \text{ and } Y := \mathbb{C} \setminus \{\pm i\}.$$

We also have $X = \overline{\pi(Z)}$ where $\pi(x, y, t) = (x, y)$ and Z is the zero locus (which is irreducible) of

$$F(x, y, t) = \begin{bmatrix} (1+t^2)x - (1-t^2) \\ (1+t^2)y - 2t \end{bmatrix}.$$

With this setup, we utilize a *pseudowitness set* [9, 10] for X which is the quadruple $\{F, \pi, L, W\}$ where L defines a linear subspace of codimension equal to the dimension of Z and W is the intersection of Z with this linear subspace defined by L . In this case, the linear polynomials L are constructed so that it has exactly $\dim X$ general linear polynomials in the image space of π , i.e., intersect X in $\deg X$ many points, while the remaining linear polynomials are general. In particular, $\pi(W)$ consists of exactly $\deg X$ distinct points. As with traditional witness sets, one can sample and perform membership tests on X [10].

The key here is that once a single sufficiently general point is known on X , other points on X can be computed as well. In fact, these other points can be forced to live in a fixed general linear subspace of codimension at most $\dim X$ thereby simplifying the future computations since one can work intrinsically on this linear subspace. If the intersection of the linear subspace and X is positive dimensional, then it is also a variety and arbitrarily many points can be sampled from this variety. If the intersection is zero-dimensional, it consists of exactly $\deg X$ points which, after computing one, random monodromy loops [19] could be used to try to compute the other points. The trace test [20] provides a stopping criterion for deciding when exactly $\deg X$ points have been computed.

Clearly, any polynomial which vanishes on X must also vanish on a finite subset of X . Although we will not delve too deep into the theory here, one can recover the invariants of X from a general linear subspace section of X when X is an *arithmetically Cohen-Macaulay* scheme (see [18, Chap. 1]). Nonetheless, since our current focus is on developing a list of potential places of where to look to focus further representation theoretic computations, we can consider all varieties and not just the arithmetically Cohen-Macaulay ones. Of course, this is at the expense of bounds rather than equality as demonstrated in the following example.

Example 2.0.2. Consider the following varieties in \mathbb{P}^3 :

$$X_1 := \{(s^3, s^2t, st^2, t^3) \mid (s, t) \in \mathbb{P}^1\} \text{ and } X_2 := \{(s^4, s^3t, st^3, t^4) \mid (s, t) \in \mathbb{P}^1\}.$$

It is easy to verify that

- $\dim X_1 = 1$, $\deg X_1 = 3$, and $I(X_1)$ is generated by three quadratics;
- $\dim X_2 = 1$, $\deg X_2 = 4$, and $I(X_2)$ is generated by a quadratic and three cubics.

Let $Y_i = X_i \cap \mathcal{H}$ be the set of $\deg X_i$ points where \mathcal{H} is the hyperplane defined by the vanishing of $\ell(x) = x_0 + 2x_1 + 3x_2 + 5x_3$. If we consider $Y_i \subset \mathcal{H}$, then

- $I(Y_1)$ is generated by three quadratics;
- $I(Y_2)$ is generated by two quadratics.

To summarize, X_1 is the twisted cubic curve in \mathbb{P}^3 which is arithmetically Cohen-Macaulay so that, for example, the dimension of $I_d(X_1)$ can be determined from $I_d(Y_1)$. However, X_2 is not arithmetically Cohen-Macaulay which, in this case, can be observed since $2 = \dim I_2(Y_2) > \dim I_2(X_2) = 1$. Even though one should only expect $d_{X_2} \geq d_{Y_2}$, we do have equality in this case, namely $d_{X_2} = d_{Y_2} = 2$.

Once we have decided on our first finite set to consider, the next task is to compute polynomials that vanish on this finite set, e.g., via [7]. This first computation alone provides some restrictions on which polynomials can be in $I(X)$. Nevertheless, we also consider what happens when we add new points to our finite set. For a particular degree, there are two possible choices: either the originally computed polynomials will vanish at the new points or the dimension of the set of polynomials that vanish at all the points will decrease. In the former case, we can then move on to searching for higher degree polynomials not generated by these polynomials. In the latter case, we continue adding new points. If no polynomials of a particular degree, say d , vanish on some finite set, then we know that $\dim I_d(X) = 0$ and $d_X > d$. Thus, we try again by considering polynomials of degree $d + 1$.

Variations of this approach can be to consider sampling points from the intersection of X with linear spaces of increasing dimension to see how the dimension of the vanishing polynomials change as less restrictions are placed on the sample points. The key in the end is to control the growth of the dimension of the space of polynomials under consideration since this can become unwieldy quickly. In particular, this method is practical for varieties X of low codimension since we can work implicitly on linear spaces of low dimension. Also, as the degrees of the polynomials under consideration increase, preconditioning becomes essential to perform reliable computations, e.g., see [7].

When the codimension is one, X is a hypersurface so that the degree of X is equal to the degree of the polynomial defining X . In this case, one can simply compute a pseudowitness set to compute its degree rather than use this interpolation based approach. For example, such an approach was used in [2] for computing the degree of implicitly defined hypersurfaces, which arise as the algebraic boundaries of Hilbert's sums of squares cones of degree 38,475 and 83,200.

Example 2.0.3. In Example 2.0.2, we considered finite sets obtained by intersecting the curves with a particular hyperplane. We now use this information to limit our focus when we add other points to our finite set. In four variables, there is a ten-dimensional space of homogeneous polynomials of degree 2, but with our previously computed information, this has already been reduced to a seven and six dimensional space for X_1 and X_2 , respectively. More specifically, the four dimensional space arising from the linear polynomial $\ell(x)$ along with the three and two dimensional spaces, respectively, from $I_2(Y_1)$ and $I_2(Y_2)$, namely

$$\begin{aligned} \bullet \quad I_2(X_1) &\subset \text{span} \left\{ \begin{array}{l} x_0\ell(x), x_1\ell(x), x_1x_2 + 2x_1x_3 + 3x_2x_3 + 5x_3^2, \\ x_2\ell(x), x_3\ell(x), x_2^2 - x_1x_3, x_1^2 - x_1x_3 - x_2x_3 - 10x_3^2 \end{array} \right\}; \\ \bullet \quad I_2(X_2) &\subset \text{span} \left\{ \begin{array}{l} x_0\ell(x), x_1\ell(x), x_1x_2 + 2x_1x_3 + 3x_2x_3 + 5x_3^2, \\ x_2\ell(x), x_3\ell(x), x_1^2 - x_2^2 + 11x_1x_3 + 2x_2x_3 + 20x_3^2 \end{array} \right\}. \end{aligned}$$

By selecting additional random points, one indeed finds $\dim I_2(X_1) = 3$ and $\dim I_2(X_2) = 1$.

This procedure develops ideas on where the degrees d_j in which generators of the ideal appear. The next step is to conclusively determine the linear subspace of the space of polynomials of degrees d_j that are in the ideal. For this, one uses representation theory as we describe next.

3. POLYNOMIALS ON THE SPACE OF BILINEAR MAPS

3.1. Tensors. In order to explain the polynomials it will be useful to work more invariantly, so instead of $\mathbb{C}^{\mathbf{a}}, \mathbb{C}^{\mathbf{b}}$ etc., we write A, B etc. for complex vector spaces of dimensions \mathbf{a}, \mathbf{b} etc.. It will also be useful to introduce the language of *tensors*. A bilinear map $A^* \times B^* \rightarrow C$ may also be viewed as a tri-linear map $A^* \times B^* \times C^* \rightarrow \mathbb{C}$, as well as in numerous other ways. To avoid prejudicing ourselves, we simply write $T \in A \otimes B \otimes C$ for any of these manifestations and call T a *tensor*. Just as we may view a linear map as a matrix after fixing bases, such T may be viewed as a three-dimensional matrix after fixing bases. Note that $A \otimes B \otimes C$, the set of all such tensors, is a vector space of dimension \mathbf{abc} . More generally, given vector spaces A_1, \dots, A_k , one can define the space of tensors $A_1 \otimes \dots \otimes A_k$. There is a natural map $A_1 \otimes \dots \otimes A_k \times B_1 \otimes \dots \otimes B_l \rightarrow A_1 \otimes \dots \otimes A_k \otimes B_1 \otimes \dots \otimes B_l$, $(f, g) \mapsto f \otimes g$, where $f \otimes g(\alpha_1, \dots, \alpha_k, \beta_1, \dots, \beta_l) := f(\alpha_1, \dots, \alpha_k)g(\beta_1, \dots, \beta_l)$.

3.2. Remarks on the theory. We briefly review the representation theory underlying in the algorithm. For more details, see [15, Chap. 6]. Let $S^d(A \otimes B \otimes C)^*$ denote the vector space of homogeneous polynomials of degree d on $A \otimes B \otimes C$. The variety $\sigma_{r;\mathbf{a},\mathbf{b},\mathbf{c}}$ is mapped to itself under changes of bases in each of the vector spaces and thus if we have one equation, we can obtain many more by changing bases. That is, let $GL(A)$ denote the set of invertible linear maps $A \rightarrow A$ and similarly for B, C . The group $G := GL(A) \times GL(B) \times GL(C)$ acts on $A \otimes B \otimes C$ by $(g_A, g_B, g_C) \cdot (\sum_i a_i \otimes b_i \otimes c_i) = \sum_i g_A a_i \otimes g_B b_i \otimes g_C c_i$, and $GL(V)$ acts on $S^d V^*$ by $g \cdot P(x) = P(g^{-1} \cdot x)$. Letting $V = A \otimes B \otimes C$ and noting $G \subset GL(V)$, we have a G -action on $S^d(A \otimes B \otimes C)^*$. If $P \in I(\sigma_r)$, then $g \cdot P \in I(\sigma_r)$ for all $g \in G$. Since ideals are in particular vector spaces, the linear span of the orbit of P in $S^d(A \otimes B \otimes C)^*$ will be in $I(\sigma_r)$.

We will use the action of the group G to organize our calculations. A group G is said to *act* on a vector space V if there is a group homomorphism $\rho : G \rightarrow GL(V)$. Then V is called a G -*module*. The G -module V is said to be *irreducible* if there is no nontrivial subspace of V invariant under the action of G . The irreducible polynomial $GL(V)$ modules are indexed by partitions $\pi = (p_1, \dots, p_{\mathbf{v}})$, where $p_1 \geq \dots \geq p_{\mathbf{v}} \geq 0$. We write $|\pi| = p_1 + \dots + p_{\mathbf{v}}$, and we say π is a partition of $|\pi|$. Let $S_{\pi} V$ denote the corresponding irreducible $GL(V)$ -module. It occurs in $V^{\otimes |\pi|}$ and no other tensor power, however not uniquely - there is a vector space's worth of realizations except in the cases $\pi = (d)$ or $\pi = (1, \dots, 1)$. The irreducible $GL(A) \times GL(B) \times GL(C)$ -modules are all of the form $V_A \otimes V_B \otimes V_C$ where V_A is an irreducible $GL(A)$ -module etc.. For $G = GL(A) \times GL(B) \times GL(C)$, every G -module decomposes into a direct sum of irreducible submodules. This decomposition is not unique in general, but the *isotypic* decomposition, where all isomorphic modules are grouped together, is.

We are interested in the homogeneous polynomials of degree say d on $A \otimes B \otimes C$, denoted $S^d(A \otimes B \otimes C)^*$. Via *polarization*, a polynomial may be considered as a symmetric tensor so $S^d(A \otimes B \otimes C)^* \subset (A \otimes B \otimes C)^{* \otimes d} \simeq A^{* \otimes d} \otimes B^{* \otimes d} \otimes C^{* \otimes d}$. Thus, the isomorphism types of irreducible G -modules in $S^d(A \otimes B \otimes C)^*$ are described by triples (π, μ, ν) of partitions of d whose number of parts $\ell(\pi) \leq \mathbf{a}$ etc.. Let $k_{\pi, \mu, \nu}$ denote the multiplicity of $S_{\pi} A^* \otimes S_{\mu} B^* \otimes S_{\nu} C^*$ in $S^d(A \otimes B \otimes C)^*$, that is, the dimension of the space of realizations of $S_{\pi} A^* \otimes S_{\mu} B^* \otimes S_{\nu} C^*$ in $S^d(A \otimes B \otimes C)^*$. The integers $k_{\pi, \mu, \nu}$ are called *Kronecker coefficients* and can be computed combinatorially. The programs SCHUR or SAGE or several other ones will compute them for you in

small cases. We used a program written by Harm Derksen, which is based on characters of the symmetric group.

There is a simple formula for $\dim S_\pi A^*$, namely

$$\dim S_{(p_1, \dots, p_a)} A^* = \prod_{1 \leq i < j \leq k} \frac{p_i - p_j + j - i}{j - i}$$

see, e.g., [6, Thm 6.3]. We will be interested in cases where the dimension is small.

Let \mathfrak{S}_d denote the group of permutations of d elements. If $\mathbf{a} = \mathbf{b} = \mathbf{c}$, then $\sigma_{r, \mathbf{a}, \mathbf{a}, \mathbf{a}}$ is also invariant under the \mathfrak{S}_3 -action permuting the vector spaces. Thus anytime $S_{\pi_1} A^* \otimes S_{\pi_2} B^* \otimes S_{\pi_3} C^*$ is in the ideal of σ_r , the module $S_{\pi_{\sigma(1)}} A^* \otimes S_{\pi_{\sigma(2)}} B^* \otimes S_{\pi_{\sigma(3)}} C^*$ will be as well, for any $\sigma \in \mathfrak{S}_3$.

3.3. First algorithm: to obtain a sample collection of polynomials. What follows is an algorithm to compute a basis of highest weight vectors for each isotypic component in $S^d(A \otimes B \otimes C)^*$. Once one has these, for each isotypic component, one can test if there are modules in the ideal of $\sigma_{r; \mathbf{a}, \mathbf{b}, \mathbf{c}}$ (or any G -variety for that matter) by sampling random points on $\sigma_{r; \mathbf{a}, \mathbf{b}, \mathbf{c}}$ as described in the second algorithm. For $\sigma \in \mathfrak{S}_d$, we write $\sigma(v_1 \otimes \dots \otimes v_d) := v_{\sigma(1)} \otimes \dots \otimes v_{\sigma(d)}$. Once and for all fix bases $a^1, \dots, a^{\mathbf{a}}$ of A^* , and similarly for B, C . Let $\pi = (p_1, \dots, p_\ell, 0, \dots, 0)$ be a partition as above. Write $\pi = (p_1, \dots, p_\ell)$ and $\ell(\pi) = \ell$. Define $F_{A, \pi} \in A^{*\otimes d}$ by

$$F_{A, \pi} := (a^1)^{\otimes(p_1-p_2)} \otimes (a^1 \wedge a^2)^{\otimes(p_2-p_3)} \otimes \dots \otimes (a^1 \wedge \dots \wedge a^f)^{\otimes(p_f-p_{f-1})}.$$

Here $v_1 \wedge \dots \wedge v_k := \frac{1}{k!} \sum_{\sigma \in \mathfrak{S}_k} \text{sgn}(\sigma) \sigma(v_1 \otimes \dots \otimes v_k)$.

Input: Degree d and partitions π, μ, ν of d .

Output: A basis P of the highest weight space vector of the isotypic component of $S_\pi A^* \otimes S_\mu B^* \otimes S_\nu C^*$ in $S^d(A \otimes B \otimes C)^*$.

- 1: Use your favorite method to compute $k_{\pi, \mu, \nu}$.
- 2: Set $k = 0$.
- 3: **while** $k < k_{\pi, \mu, \nu}$ **do**
- 4: **repeat**
- 5: Choose permutations $\tau_1, \tau_2 \in \mathfrak{S}_d$.
- 6: Define

$$F_{\pi, \mu, \nu}^{\tau_1, \tau_2} := F_{A, \pi} \otimes (\tau_1 \cdot F_{B, \mu}) \otimes (\tau_2 \cdot F_{C, \nu}) \in A^{*\otimes d} \otimes B^{*\otimes d} \otimes C^{*\otimes d},$$

rearrange the factors so it is expressed as an element of $(A \otimes B \otimes C)^{*\otimes d}$, and symmetrize to get

$$\begin{aligned} P_{\pi, \mu, \nu}^{\tau_1, \tau_2} &:= \sum_{\sigma \in \mathfrak{S}_d} \sigma \cdot F_{\pi, \mu, \nu}^{\tau_1, \tau_2} \\ &= \sum_{\sigma \in \mathfrak{S}_d} (\sigma \cdot F_{A, \pi}) \otimes (\sigma \cdot \tau_1 \cdot F_{B, \mu}) \otimes (\sigma \cdot \tau_2 \cdot F_{C, \nu}) \in A^{*\otimes d} \otimes B^{*\otimes d} \otimes C^{*\otimes d} \end{aligned}$$

where recall $\sigma \cdot (a^1 \otimes \dots \otimes a^d) := a^{\sigma(1)} \otimes \dots \otimes a^{\sigma(d)}$.

- 7: **until** $P_{\pi, \mu, \nu}^{\tau_1, \tau_2}$ is linearly independent of P_1, \dots, P_{k-1} .
- 8: Increase $k = k + 1$.
- 9: Set $P_k = P_{\pi, \mu, \nu}^{\tau_1, \tau_2}$.

10: **end while**

3.4. Examples.

3.4.1. $d = 2$, $(\pi, \mu, \nu) = ((2), (1, 1), (1, 1))$. Here $k_{(2),(1,1),(1,1)} = 1$, so we are looking for a single polynomial. We have $F_{A,(2)} = (a^1)^2$, $F_{B,(1,1)} = b^1 \wedge b^2$ and $F_{C,(1,1)} = c^1 \wedge c^2$. Try $\tau_1 = \tau_2 = Id$, then

$$\begin{aligned} F_{(2),(1,1),(1,1)}^{Id,Id} &= (a^1 \otimes a^1) \otimes (b^1 \otimes b^2 - b^2 \otimes b^1) \otimes (c^1 \otimes c^2 - c^2 \otimes c^1) \\ &= (a^1 \otimes b^1 \otimes c^1) \otimes (a^1 \otimes b^2 \otimes c^2) - (a^1 \otimes b^1 \otimes c^2) \otimes (a^1 \otimes b^2 \otimes c^1) \\ &\quad - (a^1 \otimes b^2 \otimes c^1) \otimes (a^1 \otimes b^1 \otimes c^2) + (a^1 \otimes b^2 \otimes c^2) \otimes (a^1 \otimes b^1 \otimes c^1) \end{aligned}$$

Thus

$$P_{(2),(1,1),(1,1)}^{Id,Id}(x^{ijk} a_i \otimes b_j \otimes c_k) = 2x^{111}x^{122} - 2x^{112}x^{121}$$

Here, and throughout, repeated indices are to be summed over. Note that if $T = x^{ijk} a_i \otimes b_j \otimes c_k$ has rank one, then $P_{(2),(1,1),(1,1)}^{Id,Id}(T) = 0$, but P will evaluate to be nonzero on a general rank two tensor.

3.4.2. $d = 3$, $(\pi, \mu, \nu) = ((2, 1), (2, 1), (2, 1))$. Here $k_{(2,1),(2,1),(2,1)} = 1$, so again we are looking for a single polynomial. We have $F_{A,(2,1)} = a^1 \otimes (a^1 \wedge a^2)$, and similarly for B, C . Try $\tau_1 = \tau_2 = Id$, then

$$\begin{aligned} F_{(2,1),(2,1),(2,1)}^{Id,Id} &= (a^1 \otimes a^1 \otimes a^2 - a^1 \otimes a^2 \otimes a^1) \otimes (b^1 \otimes b^1 \otimes b^2 - b^1 \otimes b^2 \otimes b^1) \otimes (c^1 \otimes c^1 \otimes c^2 - c^1 \otimes c^2 \otimes c^1) \\ &= (a^1 \otimes b^1 \otimes c^1) \otimes (a^1 \otimes b^1 \otimes c^1) \otimes (a^2 \otimes b^2 \otimes c^2) - (a^1 \otimes b^1 \otimes c^1) \otimes (a^1 \otimes b^1 \otimes c^2) \otimes (a^2 \otimes b^2 \otimes c^1) \\ &\quad - (a^1 \otimes b^1 \otimes c^1) \otimes (a^1 \otimes b^2 \otimes c^1) \otimes (a^2 \otimes b^1 \otimes c^2) + (a^1 \otimes b^1 \otimes c^1) \otimes (a^1 \otimes b^2 \otimes c^2) \otimes (a^2 \otimes b^1 \otimes c^1) \\ &\quad - (a^1 \otimes b^1 \otimes c^1) \otimes (a^2 \otimes b^1 \otimes c^1) \otimes (a^1 \otimes b^2 \otimes c^2) + (a^1 \otimes b^1 \otimes c^1) \otimes (a^2 \otimes b^1 \otimes c^2) \otimes (a^1 \otimes b^2 \otimes c^1) \\ &\quad + (a^1 \otimes b^1 \otimes c^1) \otimes (a^2 \otimes b^2 \otimes c^1) \otimes (a^1 \otimes b^1 \otimes c^2) - (a^1 \otimes b^1 \otimes c^1) \otimes (a^2 \otimes b^2 \otimes c^2) \otimes (a^1 \otimes b^1 \otimes c^1) \end{aligned}$$

Thus $P_{(2,1),(2,1),(2,1)}^{Id,Id} \equiv 0$ so we need to try different τ_1, τ_2 . Take $\tau_1 = Id$ and $\tau_2 = (12)$. Then

$$\begin{aligned} F_{(2,1),(2,1),(2,1)}^{Id,(1,2)} &= (a^1 \otimes a^1 \otimes a^2 - a^1 \otimes a^2 \otimes a^1) \otimes (b^1 \otimes b^1 \otimes b^2 - b^1 \otimes b^2 \otimes b^1) \otimes (c^1 \otimes c^1 \otimes c^2 - c^2 \otimes c^1 \otimes c^1) \\ &= (a^1 \otimes b^1 \otimes c^1) \otimes (a^1 \otimes b^1 \otimes c^1) \otimes (a^2 \otimes b^2 \otimes c^2) - (a^1 \otimes b^1 \otimes c^2) \otimes (a^1 \otimes b^1 \otimes c^1) \otimes (a^2 \otimes b^2 \otimes c^1) \\ &\quad - (a^1 \otimes b^1 \otimes c^1) \otimes (a^1 \otimes b^2 \otimes c^1) \otimes (a^2 \otimes b^1 \otimes c^2) + (a^1 \otimes b^1 \otimes c^2) \otimes (a^1 \otimes b^2 \otimes c^1) \otimes (a^2 \otimes b^1 \otimes c^1) \\ &\quad - (a^1 \otimes b^1 \otimes c^1) \otimes (a^2 \otimes b^1 \otimes c^1) \otimes (a^1 \otimes b^2 \otimes c^2) + (a^1 \otimes b^1 \otimes c^2) \otimes (a^2 \otimes b^1 \otimes c^1) \otimes (a^1 \otimes b^2 \otimes c^1) \\ &\quad + (a^1 \otimes b^1 \otimes c^1) \otimes (a^2 \otimes b^2 \otimes c^1) \otimes (a^1 \otimes b^1 \otimes c^2) - (a^1 \otimes b^1 \otimes c^2) \otimes (a^2 \otimes b^2 \otimes c^1) \otimes (a^1 \otimes b^1 \otimes c^1). \end{aligned}$$

$$\begin{aligned} \text{Thus } P_{(2,1),(2,1),(2,1)}^{Id,(1,2)} &\left(\sum_{i,j,k=1}^2 x^{ijk} a_i \otimes b_j \otimes c_k \right) = \\ &x^{111}x^{111}x^{222} + 2x^{112}x^{121}x^{211} - x^{111}x^{121}x^{212} - x^{111}x^{211}x^{122} - x^{111}x^{112}x^{221}. \end{aligned}$$

Note that if T has rank one, then $P_{(2,1),(2,1),(2,1)}^{Id,(1,2)}(T) = 0$, but P will evaluate to be nonzero on a general rank two tensor.

3.4.3. *Permutation pairs to avoid.* We want to avoid the case that occurs in Example 3.4.2, i.e., that $P_{\pi,\mu,\nu}^{\tau_1,\tau_2} = 0$. Although a complete classification of the cases when this happens is unknown, an easy necessary condition for $P_{\pi,\mu,\nu}^{\tau_1,\tau_2} \neq 0$ is the following [11, Lemma 7.2.7]: When we write $1, 2, \dots, d$ in a tableau columnwise starting with the longest column and we write $\tau_1(1), \tau_1(2), \dots, \tau_1(d)$ in a second tableau columnwise, and we do the same for τ_2 in a third tableau, then it is necessary that there exists no pair of numbers that lies in the same column in all three tableaux. If this occurs, we call this situation a *zero pattern*. We can choose random

permutations that avoid the zero pattern by just choosing random permutations and repicking if it contains a zero pattern.

3.4.4. $d = 20$, $(\pi, \mu, \nu) = ((5555), (5555), (5555))$. Here $k_{(5555), (5555), (5555)} = 4$. The following choices of pairs τ_1, τ_2 give 4 linearly independent polynomials.

$$\tau_1 = (\tau_1(1), \tau_1(2), \dots, \tau_1(20)) = (10, 15, 5, 9, 13, 4, 17, 14, 7, 20, 19, 11, 2, 12, 8, 3, 16, 18, 6, 1),$$

$$\tau_2 = (10, 11, 6, 2, 8, 9, 4, 20, 15, 16, 13, 18, 14, 19, 7, 5, 17, 3, 12, 1)$$

$$\tau_1 = (19, 10, 1, 5, 7, 12, 2, 13, 16, 6, 18, 9, 11, 20, 3, 17, 14, 8, 15, 4),$$

$$\tau_2 = (10, 5, 13, 6, 3, 16, 11, 1, 4, 18, 15, 17, 9, 2, 8, 12, 19, 7, 14, 20)$$

$$\tau_1 = (16, 20, 9, 13, 8, 1, 4, 19, 11, 17, 7, 2, 14, 3, 6, 5, 12, 15, 18, 10),$$

$$\tau_2 = (1, 20, 11, 19, 5, 16, 17, 2, 18, 13, 7, 12, 14, 10, 8, 15, 6, 9, 3, 4)$$

$$\tau_1 = (11, 5, 2, 1, 16, 10, 20, 3, 17, 19, 12, 18, 13, 9, 14, 4, 8, 6, 15, 7),$$

$$\tau_2 = (1, 6, 15, 13, 20, 3, 18, 11, 14, 2, 9, 5, 4, 17, 12, 8, 19, 16, 7, 10)$$

This gives rise to 4 polynomials f_1, \dots, f_4 , and the following linear combination vanishes on σ_6 : $-266054 f_1 + 421593 f_2 + 755438 f_3 + 374660 f_4$.

3.4.5. $d = 19$, $(\pi, \mu, \nu) = ((5554), (5554), (5554))$. Here $k_{(5554), (5554), (5554)} = 31$. We found 31 pairs τ_1, τ_2 that result in 31 linearly independent polynomials by choosing τ_1 and τ_2 randomly, but avoiding the zero pattern.

3.5. Second algorithm: to test on the secant variety. Once one has a basis of highest weight vectors for an isotypic component, one needs to determine which linear combinations of basis vectors vanish on σ_r . The following algorithm is standard linear algebra:

Input: The output of first algorithm for some (π, μ, ν) , i.e., a collection $P_1, \dots, P_k = P_{k\pi, \mu, \nu} \in S^d(A \otimes B \otimes C)^*$ and r , where we will test for polynomials in $I(\sigma_{r; \mathbf{a}, \mathbf{b}, \mathbf{c}})$.

Output: with probability as high as you like the component of $I(\sigma_{r; \mathbf{a}, \mathbf{b}, \mathbf{c}})$ in $S_\pi A^* \otimes S_\mu B^* \otimes S_\nu C^*$. If the component is zero, then the answer is guaranteed correct, and more generally, the algorithm can only overestimate the component if the points on σ_r are not chosen randomly enough.

- 1: Set $P = c_1 P_1 + \dots + c_k P_k$, where c_1, \dots, c_k are variables.
- 2: Chose “random” vectors

$$v_j = \sum_{i=1}^{\mathbf{a}} \sum_{k=1}^{\mathbf{b}} \sum_{l=1}^{\mathbf{c}} (\alpha_{1,j}^i a_i) \otimes (\beta_{1,j}^k b_k) \otimes (\gamma_{1,j}^l c_l) + \dots + (\alpha_{r,j}^i a_i) \otimes (\beta_{r,j}^k b_k) \otimes (\gamma_{r,j}^l c_l)$$

where the $\alpha_{\delta,j}^i, \beta_{\delta,j}^k, \gamma_{\delta,j}^l$ are “random” numbers.

- 3: Evaluate P at these k points.
 - 4: **if** there exist a solution c_1, \dots, c_k such that all the evaluations are zero **then**
 - 5: If there is a m -dimensional solution space, then with reasonable probability one has m copies of the module in the ideal.
 - 6: **else**
 - 7: No module in this isotypic component is in $I(\sigma_{r; \mathbf{a}, \mathbf{b}, \mathbf{c}})$.
 - 8: **end if**
-

Remark 3.5.1. One can parametrize σ_r and in some cases use the explicit parametrization to get an exact answer, which is what we did for $(\pi, \mu, \nu) = ((5555), (5555), (5555))$.

4. REVIEW OF THE ORIGINAL PROOF THAT $M_2 \notin \sigma_{6;4,4,4}$

The essence of the proof that the rank of M_2 is not six in [14] is as follows: there is a now standard argument due to Baur for proving lower bounds for rank by splitting a putative computation into two parts using the algebra structure on the space of matrices. The argument in [14] was to apply the same type of argument to each component of the variety consisting of subvarieties where the rank is greater than the border rank. The article [14] contained a gap in the proof that was filled in [12] but not published in JAMS because the editor was concerned the erratum was almost as long as the original article and the author did not see a way to shorten it. The gap in [14] was caused by overlooking the possibility of certain types of components, where the limiting 6-planes are not formed by points coming together but by some other unusual configuration of points. All such components of $\sigma_{6;4,4,4}$ are not known explicitly, but the correction only used qualitative aspects of how the limiting 6-plane arose. There were 3 basic cases, where any subset of 5 of the limit points are linearly independent, where there is a subset of 5 that are not, but any subset of four are, and where there is a subset of 4 that are not, but any subset of 3 are. In each of these cases, one is forced to have a limit taking place among rank one tensors in a much smaller space, which was what made the analysis tractable. The computations performed above provide an explicit polynomial vanishing on $\sigma_{6;4,4,4}$ which does not vanish at M_2 , providing a significantly shorter proof of this fact.

REFERENCES

1. D. Bates and L. Oeding, *Toward a salmon conjecture*, preprint, arXiv:1009.6181.
2. G. Blekherman, J. Hauenstein, J.C. Ottem, K. Ranestad, and B. Sturmfels, *Algebraic boundaries of Hilbert's SOS cones*, *Compositio Mathematica* **148** (2012), 1717–1735.
3. Peter Bürgisser, Michael Clausen, and M. Amin Shokrollahi, *Algebraic complexity theory*, Grundlehren der Mathematischen Wissenschaften, vol. 315, Springer-Verlag, Berlin, 1997, With the collaboration of Thomas Lickteig. MR 1440179 (99c:68002)
4. Peter Bürgisser and Christian Ikenmeyer, *Geometric Complexity Theory and Tensor Rank*, *Proceedings 43rd Annual ACM Symposium on Theory of Computing 2011* (2011), 509–518.
5. ———, *Explicit Lower Bounds via Geometric Complexity Theory*, to appear in *Proceedings 45th Annual ACM Symposium on Theory of Computing 2013*, preprint arXiv:1210.8368 (2013).
6. William Fulton and Joe Harris, *Representation theory*, Graduate Texts in Mathematics, vol. 129, Springer-Verlag, New York, 1991, A first course, Readings in Mathematics. MR MR1153249 (93a:20069)
7. Z.A. Griffin, J.D. Hauenstein, C. Peterson, and A.J. Sommese, *Numerical computation of the Hilbert function of a zero-scheme*, Available at www.math.ncsu.edu/~jdhouens/preprints.
8. J.D. Hauenstein, Y.-H. He, and D. Mehta, *Numerical analyses on moduli space of vacua*, arXiv:1210.6038, 2012.
9. J.D. Hauenstein and A.J. Sommese, *Witness sets of projections*, *Appl. Math. Comput.* **217** (2010), no. 7, 3349–3354.
10. ———, *Membership tests for images of algebraic sets by linear projections*, 2013, pp. 6809–6818.
11. Christian Ikenmeyer, *Geometric Complexity Theory, Tensor Rank, and Littlewood-Richardson Coefficients*, Ph.D. thesis, Institute of Mathematics, University of Paderborn, 2012, Online available at http://math-www.uni-paderborn.de/agpb/work/ikenmeyer_thesis.pdf.
12. J. M. Landsberg, *The border rank of the multiplication of 2×2 matrices is seven*, arXiv:math/0407224.
13. ———, *Explicit tensors of border rank at least $2n-1$* , preprint arXiv:1209.1664.
14. ———, *The border rank of the multiplication of 2×2 matrices is seven*, *J. Amer. Math. Soc.* **19** (2006), no. 2, 447–459 (electronic). MR MR2188132 (2006j:68034)
15. ———, *Tensors: geometry and applications*, Graduate Studies in Mathematics, vol. 128, American Mathematical Society, Providence, RI, 2012. MR 2865915
16. J. M. Landsberg and Laurent Manivel, *On the ideals of secant varieties of Segre varieties*, *Found. Comput. Math.* **4** (2004), no. 4, 397–422. MR MR2097214 (2005m:14101)
17. J.M. Landsberg and Giorgio Ottaviani, *New lower bounds for the border rank of matrix multiplication*, preprint, arXiv:1112.6007.

18. J.C. Migliore, *Introduction to liaison theory and deficiency modules*, Progress in Mathematics, vol. 165, Birkhäuser Boston Inc., Boston, MA, 1998.
 19. A.J. Sommese, J. Verschelde, and C.W. Wampler, *Using monodromy to decompose solution sets of polynomial systems into irreducible components*, Applications of algebraic geometry to coding theory, physics and computation (Eilat, 2001), NATO Sci. Ser. II Math. Phys. Chem., vol. 36, Kluwer Acad. Publ., Dordrecht, 2001, pp. 297–315.
 20. ———, *Symmetric functions applied to decomposing solution sets of polynomial systems*, SIAM J. Numer. Anal. **40** (2002), no. 6, 2026–2046.
 21. A.J. Sommese and C.W. Wampler, II, *The numerical solution of systems of polynomials*, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2005, Arising in engineering and science.
- E-mail address:* hauenstein@ncsu.edu, ciken@math.upb.de, jml@math.tamu.edu